

Krügel, Sebastian; Uhl, Matthias

Article — Published Version

The behavioral economics of dynamically inconsistent behavior: a critical assessment

Social Choice and Welfare

Provided in Cooperation with:

Springer Nature

Suggested Citation: Krügel, Sebastian; Uhl, Matthias (2023) : The behavioral economics of dynamically inconsistent behavior: a critical assessment, Social Choice and Welfare, ISSN 1432-217X, Springer, Berlin, Heidelberg, Vol. 61, Iss. 4, pp. 817-833, <https://doi.org/10.1007/s00355-023-01471-5>

This Version is available at:

<https://hdl.handle.net/10419/308950>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



The behavioral economics of dynamically inconsistent behavior: a critical assessment

Sebastian Krügel^{1,2} · Matthias Uhl¹

Received: 10 August 2021 / Accepted: 3 June 2023 / Published online: 14 July 2023
© The Author(s) 2023

Abstract

Preferences often change—even in short time intervals—due to either the mere passage of time (present-biased preferences) or changes in visceral or environmental conditions (state-dependent preferences). On the basis of empirical findings concerning state-dependent preferences, we critically discuss the “Aristotelian” view of unitary decision makers in economics. We illustrate that the conceptualization of preferences as “present-biased” as opposed to “state-dependent” has very different normative implications for which preferences should be considered “rational.” Empirically, however, the two concepts are very difficult to distinguish. The economist can justify any paternalistic intervention if she can conceptualize changing preferences so flexibly, and she can easily become a benevolent despot. We therefore urge for a more careful “Heraclitean” view of decision-making that accepts that a person may consist of multiple selves.

1 Introduction

People’s choices often are dynamically inconsistent. They make a plan for a future period but revise that plan when the future period arrives even though no new information occurred. In this article, we criticize the paternalistic stance that most behavioral economists tend to take toward this human peculiarity. Typically, behavioral economists discuss dynamic inconsistency in terms of a behavioral “bias.” There

✉ Sebastian Krügel
sebastian.kruegel@thi.de

Matthias Uhl
matthias.uhl@thi.de

¹ Technische Hochschule Ingolstadt, Esplanade 10, D-85049 Ingolstadt, Germany

² Technical University of Munich, Arcisstrasse 21, D-80333 Munich, Germany

are, however, two very different conceptualizations of dynamic inconsistency in the pertinent literature. First, dynamic inconsistency is conceptualized as a problem of *present-biased preferences*.¹ This conceptualization aims to capture the human tendency toward immediate gratification. Here, an individual makes a plan for the future and then gives in to a tempting impulse in the moment of consumption. It is explicitly acknowledged that some individuals are sophisticated enough to foresee their bias for the present and attempt to constrain themselves later.

Second, dynamic inconsistency is conceptualized as a problem of *projection bias* (Loewenstein et al. 2003) when preferences are state-dependent. Contrary to the idea of a present bias, where the “error” occurs at the time of consumption, the behavioral mistake in the model of projection bias occurs at the time of decision (i.e., while *planning* future consumption). Here, an individual mistakenly projects her current desires in the moment of decision to a different visceral or environmental state at the future moment of consumption in which her preferences will systematically deviate from her current ones. Unlike the literature on present bias, the literature on projection bias presupposes that the individual must be unaware of her bias.

It is crucial to note that these two conceptualizations of dynamic inconsistency come to radically different conclusions about the normative question of whether the preferences at the time of decision or at the time of consumption are the “true” ones. Proponents of present-biased preferences typically take sides with the “long-run” perspective and view the preferences at the time of decision as the “true” ones (see, e.g., Akerlof 1991; O’Donoghue and Rabin 1999, 2001). In the context of state-dependent preferences, it is the other way around. Here, the “preferences that should be relevant are those that will prevail when the consequences occur” (Read and Van Leeuwen 1998, p. 189). To justify a desired paternalistic intervention, one would only have to refer to the conceptualization that attaches the derogatory label “bias” to the disliked preference. This is particularly problematic, if—as we will argue—the empirical studies that are used to establish the behavioral bias do not test for alternative conceptualizations of dynamic inconsistencies. Until the burden of proof for intervention is established, we propose a third and more cautious conceptualization of dynamic inconsistency, namely intrapersonal conflict. It rejects the concept of “true” preferences and considers dynamically inconsistent behavior as a strategic interaction between alternating preferences—a kind of self-management. While a “bias” (whether in projection or in the present) calls for external intervention to correct the failure, rational self-management demands restraint.

So far, the idea of intrapersonal conflict and multiple selves has been discussed mainly in the context of present-biased preferences (see, e.g., Strotz 1955; Thaler and Shefrin 1981; Parfit 1984; O’Donoghue and Rabin 1999; Jamison and Wegener 2010). In this article, we will focus on an understanding of multiple selves in relation to the notion of state-dependent preferences. Prima facie, the state-dependency view of preferences could be more hospitable to an interpretation of decision makers

¹ We use the term of O’Donoghue and Rabin (1999) for the general phenomenon that people appear to overvalue immediate satisfaction. Another term for the same phenomenon is, for instance, (quasi-)hyperbolic discounting.

in terms of multiple selves, because it does not degrade a particular preference to a “non-rational” impulse. However, we would like to point out that many examples of supposedly present-biased preferences can also be understood as expressions of state-dependent preferences (e.g., ordering too much dessert in a restaurant, sleeping late in the morning, not going to the gym, etc.). The dividing line between typical examples of the two conceptions of preferences is usually not very sharp. Our discussion of intrapersonal conflict in the context of state-dependent preferences is therefore easily applicable to many “real-world” examples of present-biased preferences.

In the following section, we will first scrutinize the underlying assumptions of the different conceptualizations of preferences and discuss how they are philosophically related. In Sect. 3, we illustrate two exemplary experiments on state-dependent preferences in detail and outline what has been inferred from their results. In Sect. 4, we offer an alternative interpretation of the results from these two experiments and discuss what would be necessary to distinguish empirically between the different interpretations. In Sect. 5, we critically assess the Aristotelian view on individual decision-making with regard to state-dependent preferences at a conceptual level and illustrate that some implausible assumptions are necessary to defend the approach. In the final section, we conclude and outline the different philosophical and political implications of the two approaches presented for our understanding of self-management.

2 Paternalism with present-biased and state-dependent preferences: anything goes

According to standard economics, rational actors rank present choices according to preferences over the causal consequences of actions. Causality implies the lapse of time, and preferences may change between the times of choice and of consequence. Many years of research in economics and psychology have taught us that changes in preferences between these two points in time may be the rule rather than the exception. For instance, numerous studies on intertemporal decision-making demonstrated that people’s choices might reverse because of the mere passage of time. Goods that people plan to choose when consumption is in the distant future are often systematically different from the goods that people actually choose as the future draws nearer. People’s preferences often appear time-inconsistent. They are “present-biased” (O’Donoghue and Rabin 1999) in the sense that immediate consumption appears to be excessively overvalued.

However, present-biased preferences are not the only instance of changing tastes. Changes in visceral or environmental conditions might also affect people’s preferences. This second source of changing tastes has received increasing attention in the more recent economic and psychological literature because many of those changes follow a systematic and thus predictable pattern (see, e.g., Loewenstein 1996, 2000, for a general discussion of systematic fluctuations in tastes). For instance, people tend to prefer healthy (e.g., apples) over unhealthy (e.g., chocolate bars) food in a satiated state, but their preferences reverse in a hungry state (Read and Van Leeuwen

1998). The value ascribed to durable cold-weather items (e.g., winter jackets) is higher on a cold than on a warm day (Conlin et al. 2007), and the value attributed to certain types of cars (e.g., convertibles) is different on a sunny compared to a cloudy day (Busse et al. 2014). Such preferences are called state-dependent preferences, with different states being associated with different preferences. Whereas preferences under this concept vary with the underlying state (for instance, hunger, arousal or weather conditions), the concept of present-biased preferences captures the idea that preferences change with a mere movement along the timeline.

With changing preferences, rational actors are confronted with the problem of how to incorporate those changes into their rankings of alternatives when making their choices. Which preferences should a rational actor obey? Those at the time of the decision or those at the time of consumption when the consequences of the decision occur?

Whereas economists are usually concerned about the “pretence of knowledge” (von Hayek 1989) when it comes to social choices, this concern seems to be much less pronounced when it comes to individual decision-making. In stark contrast to the revealed-preference approach (Samuelson 1938, 1948), many economists appear to be able to tell what people’s relevant (i.e., true) preferences are even if choices differ between two points in time. Unfortunately, however, the economists’ answers to this question can be very different, depending on the underlying concept of preferences, as mentioned above. In the context of present-biased preferences, the standard belief in economics is that people’s true preferences are those at the time of the decision, because preferences at the time of consumption are distorted by temptation or a tendency for immediate gratification (see, e.g., Strotz 1955; Phelps and Pollak 1968; Pollak 1968; Laibson 1997; O’Donoghue and Rabin 1999). In the context of state-dependent preferences, on the contrary, the traditional view among economists is that the relevant preferences prevail at the time of consumption, not at the time of the decision (see, e.g., Read and Van Leeuwen 1998; Loewenstein et al. 2003).

Rational actors, who behave in accordance with the normative expectations of most economists, would thus have to take into account in their decisions whether a change in preferences between choice and consequence is caused by a movement along the time line or by a change of the person’s state. Truly rational actors may be able to do this, but to an outside observer, this is very difficult to tell. The problem gets thorny when it comes to questions of consumer sovereignty and paternalistic interventions. In light of present-biased preferences, the benevolent economist is inclined to promote people’s preferences at the time of the decision. In the case of state-dependent preferences, in contrast, the economist would side with people’s preferences at the time of consumption. However, because it is difficult for an external advisor to judge what causes changes in preferences, partisanship for one set of preferences over another is problematic and perhaps primarily ideologically driven.

Although it may seem surprising at first that partisanship for a certain set of preferences in economics would be so arbitrary, there is a unifying element between both approaches to changes in preferences: the Aristotelian view on human decision-making that prevails in economics (Moldoveanu and Stevenson 2001). People are regarded as unitary decision makers, possibly coupled with various behavioral biases. Examples of self-commitment are interpreted exclusively as strategic

interventions of rational actors against their own temptations in the context of present-biased preferences. The desire at the time of consumption is more of an impulse than a real preference. People with present-biased preferences are therefore “naive” (O’Donoghue and Rabin 1999) if they fail to suppress their future temptations due to a lack of farsightedness, and they are “sophisticated” (O’Donoghue and Rabin 1999) if they use self-commitment measures to preserve their current preferences against upcoming changes of mind. Ulysses is the ideal conception of rationality here. For people with state-dependent preferences, the opposite is true. Self-commitment measures are out of the question. People with state-dependent preferences are sophisticated if they abandon their current preferences in favor of their future ones and somewhat naive if they fail to please their future preferences.

Clearly, if the usual assumption of unitary decision makers is complemented with the attribution of behavioral biases, the Aristotelian view makes paternalistic interventions into the self-management of people appear attractive, because each person consists of only one true rationale (or *telos*) that is to be promoted. However, since the idea of a person’s true rationale differs between present-biased and state-dependent preferences, almost any paternalistic intervention is justifiable. In many real-life cases, the underlying reason for a preference change is hardly detectable by external advisors.

An alternative perspective that does not sanctify the assumption of a unitary decision maker and instead takes the notion of multiple selves and intrapersonal conflict seriously would likely lead to a more liberal and less invasive stance towards a person’s self-management. In the present article, we mainly discuss intrapersonal conflict in relation to the notion of state-dependent preferences because this conceptualization of preferences does not degrade a particular desire to a “non-rational” impulse. However, if the Aristotelian view is upheld, this implies that the preference at the time of consumption is the only preference that counts. According to this view, people with state-dependent preferences whose choices are not fully in line with their own future preferences must be mistaken in predicting them (see, e.g., Read and Van Leeuwen 1998; Loewenstein et al. 2003; Conlin et al. 2007; Sayette et al. 2008; Simonsohn 2010; Busse et al. 2014; Buchheim and Kolaska 2017). It is not an intrapersonal conflict between multiple selves that is observed, but a unitary decision maker’s inability to properly foresee what she wants.

3 State-dependent preferences: empirical findings and interpretations

For many real-life decisions, the temporal separation between choice and consequence amounts to several days, weeks or even months and longer. This is the case when we preorder lunch in the cafeteria for the next day, buy groceries for the upcoming weekend, order goods via the Internet or a catalog or book a vacation trip for the upcoming summer. Choices with a marked time lag between the decision and consumption are sometimes termed advance choices and have been frequently used to study state-dependent preferences. The greater time lag as such, however, is not of particular interest, but only a means to an end. The

crucial element in those studies is that the preferences—i.e., the person's *states*—differ between the two points in time.

In laboratory or field experiments, state-dependent preferences were investigated in various situations involving, for instance, hunger (Read and Van Leeuwen 1998), cigarette craving (Sayette et al. 2008), sexual arousal (Ariely and Loewenstein 2006) or pain (Read and Loewenstein 1999). Most of these studies employed a very similar procedure. The participants made an advance choice in period 1 for a predefined future period 2 with corresponding state s_2 . In period 1, some of the participants were in the same state that they occupied in period 2 (i.e., $s_1 = s_2$), while other participants were in a different state (i.e., $s'_1 \neq s_2$). The participants were led to believe that their choice in period 1 would count but when period 2 arrived, they could in fact “re-make” their choice. The typical results of those studies were that the participants made systematically different advance choices in state s'_1 compared to in state s_1 and revised their choices more often in period 2 if the advance choice was made in state s'_1 .

Consider, for instance, the classical study of Read and Van Leeuwen (1998). They conducted a field experiment in which office workers made an advance choice between healthy (e.g., apples) and unhealthy (e.g., chocolate bars) snacks, which they received at a designated time one week later. The experimenters used a 2×2 between-subjects design. First, they varied the state of hunger or satiation at the time of the decision, i.e., when making the advance choice that unfolded its consequence only one week later. Second, they varied the state of hunger or satiation at the time of consumption, i.e., when the consequence of the advance choice unfolded. Participants were then randomly allocated to one of the emerging four treatments. In two treatments, participants made an advance choice in a state identical to the state when they would actually receive the food: Currently hungry (satiated) subjects made an advance choice for a hungry (satiated) state one week later. In the two remaining treatments, participants made an advance choice in a state that was different to the state when they would actually receive the food: Currently hungry (satiated) subjects made an advance choice for a satiated (hungry) state one week later.

Participants, however, were then deceived by the experimenters. Directly before participants would receive their chosen snack at the designated time one week later, they were taken by surprise and asked to make a new choice for their immediate consumption. This was justified by the experimenters pretending that they had lost the list in which the participants' advance choices had been registered. Read and Van Leeuwen (1998) found that advance choices were indeed influenced by anticipated future hunger levels as well as current hunger levels. People who expected to be hungry the next week chose unhealthy snacks more often than people who expected to be satiated. The same was true for immediate choices. People who were hungry at the time of the immediate choice chose unhealthy snacks more often than those who were satiated. Consequently, people who had made their advance choice in a state that was different from the one in which they made their immediate choice revised their advance choice systematically more often when making their immediate choice. From this, Read and Van Leeuwen (1998) concluded that people erroneously projected their current preferences onto the future.

Consider another exemplary experiment. In a study on projection bias in smokers, Sayette et al. (2008) had smoking participants make advance choices for a later state for which they knew they would be nicotine deprived for at least 12 h, i.e., when being in a craving state. Nicotine deprivation would be assured by taking breath samples to determine participants' carbon monoxide levels. When making their advance choice, participants were asked to state the minimum amount of money that they would have to get to delay smoking for another five minutes in the next session when being in the craving state. This minimum amount was elicited in an incentive compatible way, meaning that participants were monetarily incentivized to reveal the smallest acceptable amount of money honestly. Crucially, some participants were randomly assigned to a condition in which they made their advance choice in a state in which they were already nicotine deprived for at least 12 h, i.e., in a craving state. Other participants made their advance choice in a state in which they had smoked regularly, i.e., in a satisfied state. During the second session, in which all participants were in a craving state, participants were offered a surprise possibility to restate their minimum amount for which they would be willing to delay smoking for another five minutes "right now."

Sayette et al. (2008) found that the revisions of craving smokers in the second session were stronger for those who had made their advance choice in a satisfied state as opposed to those who had made their advance choice in a craving state. From this, the authors conclude that satisfied smokers underpredicted the value that they would place on smoking in the craving state, because they would falsely project their current preferences to this different state.

Field data studies on state-dependent preferences are not as common, but they exist as well. State-dependent preferences often change with visceral factors (like hunger or arousal) that are difficult to observe. These changes may, however, sometimes be mediated by environmental conditions that can be better observed. Most of the field data studies, therefore, have utilized certain weather data to show that weather conditions systematically influence people's advance choices in a way that is inconsistent with expected utility theory. For instance, Conlin et al. (2007) investigated catalog orders of cold-weather items and found that the colder the weather on the order date, the more likely a return of that item was once it was received. Buchheim and Kolaska (2017) analyzed advance ticket sales for an outdoor movie theater and found that good weather (i.e., using sunshine duration at the time of sale as criterion) increased advance ticket sales even though the weather at the time of the purchase did not predict the weather on the day of the movie. Busse et al. (2014) studied a large data set of vehicle transactions and found that buying decisions for convertibles or four-wheel drives were affected considerably by the weather conditions on the day of the purchase. Lastly, Simonsohn (2010) even found that certain weather conditions (i.e., cloud cover) on the day of the visit to an academically demanding university affected the enrollment decisions of prospective students regarding this university.

Clearly, all of these studies demonstrated that people display state-dependent *behavior*. That is, people's advance choices appear to be overly influenced by their current preferences (or states). Based on the assumption in economics that the relevant preferences prevail at the time of consumption, a widely accepted interpretation

of the observed behavior is that people systematically mispredict their future preferences. People wish to please their future preferences in a different state, but unfortunately, they are biased towards their current wants. This misprediction is sometimes referred to as an *empathy gap* (e.g., Loewenstein 1996, 2000; Loewenstein and Schkade 1999) or *projection bias* (e.g., Loewenstein et al. 2003; Conlin et al. 2007; Busse et al. 2014). The latter term underlines the view that current preferences are illegitimately projected to another state for which a decision is to be made. People seem to understand the direction in which their preferences will change but systematically underestimate the magnitude of this change (Loewenstein et al. 2003).

In the framework of state-dependent preferences, projection bias can be introduced in the following way (see Loewenstein et al. 2003, who first formalized the idea that people with state-dependent preferences systematically mispredict their future preferences). Suppose a person's instantaneous utility of consumption in period t is given by $u(c_t, s_t)$, where c_t is the person's consumption in period t and s_t is the person's state in period t capturing her preferences. Further suppose that this person is currently in period 1 with corresponding state s_1 and is trying to predict her future instantaneous utility from consuming c_2 in period 2 with corresponding state s_2 (where $s_1 \neq s_2$). This prediction is denoted $\tilde{u}(c_2, s_2 | s_1)$. If this person has no projection bias, she will predict her future utility correctly. That is, her predicted utility will equal her true utility: $\tilde{u}(c_2, s_2 | s_1) = u(c_2, s_2)$. If, on the other hand, this person is exposed to a projection bias, as described in Loewenstein et al. (2003), she will understand the qualitative direction of the taste change but underestimate its magnitude. That is, her predicted utility will be somewhere in between her true future utility and her utility given her current state: $\tilde{u}(c_2, s_2 | s_1) = (1 - \alpha)u(c_2, s_2) + \alpha u(c_2, s_1)$, with $\alpha \in [0, 1]$. A person with projection bias is a person with $\alpha > 0$, with the bias increasing with higher values of α . Since a person with projection bias misperceives her future utility, such an individual may exhibit dynamic inconsistency even in the absence of present-biased preferences: she may make a systematically different choice in period 1 (for period 2) than if she were asked again in period 2.

Importantly, the projection-bias interpretation of state-dependent behavior is based entirely on mistaken beliefs. A person with projection bias behaves exactly as a rational person does with the one exception that she mispredicts her future preferences (Loewenstein et al. 2003). Rationality here means maximizing intertemporal utility.

4 State-dependent behavior: an alternative interpretation

Our main caveat with the interpretations of the empirical results is that the participants' intentions at the time of making their advance choices remain unknown. All that is observed is state-dependent behavior. The inference that this behavior results from systematic mispredictions of future preferences is based solely on the assumption that the relevant preferences prevail at the time of consumption and, more importantly, that people share this view.

An alternative but rarely considered interpretation of the same behavior is that people do not mispredict their future preferences in a different state but rather disagree with them and therefore try to impose a “better” judgment on a later alter ego. A person may generally disapprove of candy in a satiated state but have a less negative attitude towards sweets in a hungry state (independent of the time of consumption). Likewise, a person may spend parts of her limited budget on a certain durable good in one state, whereas she would want to use the same money for a different durable good in another state, simply because her overall valuations diverge in both states. For instance, suppose a person can allocate parts of her budget to either a handcrafted surfboard or a snowboard. Suppose it is currently summer, but since the surfboard is custom-made and making it will take a few months, it cannot be used before the next summer. The person may still order the surfboard, knowing that she would decide differently during winter. The same argument may apply when ordering or returning warm winter jackets. Note that this does not imply that the self in the warm wants to inflict suffering on the other self in the cold. Economics is about decision-making under budget constraints. The two selves may simply disagree on how to allocate their limited budget.

If people are aware of this disagreement and insist on their valuations at the time of the decision, they will attempt to impose a “better” judgment at the expense of a later alter ego. This imposition can be understood as a *sympathy gap* or *conflict of selves*. Whereas the empathy gap is an ineffective attempt to please preferences in a different state, the sympathy gap is an attempt to fight them. No previous study could potentially rule out the latter explanation. To the best of our knowledge, none of the previous studies on projection bias did even attempt to do that.

Let us reconsider the study of Read and van Leeuwen (1998), for instance. People revised their choices in the second week more often when their advanced choices one week before were made in a different state. If the only intention of a person with state-dependent preferences is to meet her future preferences in a different state as accurately as possible, the stronger revision would indeed represent a misprediction. However, it is far from clear whether a participant who chooses a healthy food in a satiated state to be received in a later hungry state wants to please this other self’s preferences as accurately as possible. It is well conceivable that the satiated self finds the hungry self’s urge to indulge undesirable and is grateful for the opportunity to commit the hungry self to its advance choice. The satiated self is then betrayed by the experimenters who surprisingly give the hungry self the option to overrule the other self’s attempt to commit to a healthy snack. The experimenters simply do not know whether people mispredicted their systematically deviating preferences in a different state or whether they predicted them with current disapproval. While self-commitment is well acknowledged in the context of present-biased preferences, it is oddly out of the picture when state-dependent preferences are discussed. To the contrary, deviating advance choices as an attempt to promote one’s interests at the time of decision, the hallmark of sophistication in the realm of present-biased preferences, is considered a projection bias in the realm of state-dependent preferences.

The same argument applies to the study by Sayette et al. (2008). Here, smokers in a currently satisfied state asked for lower compensations for delaying smoking in a later craving state in their advance choices than smokers in a currently craving state

did in their advance choices. It is far from clear that this represents a misprediction of their urge to smoke in a craving state. It could have well been the expression of a disapproval with one's addictive tendencies in the condition of nicotine deprivation. Interestingly, Sayette et al. (2008) also measured participants self-reported urge to smoke. Participants were asked to report their current urge to smoke in each session as well as their anticipated urge during the second session. Sayette et al. (2008) found that participants who were in a satisfied state did not do significantly worse in predicting their future urge to smoke during the second high-craving session than participants who were already in a high-craving state during the first session. The observation that participants in a satiated state still demanded lower compensations for delaying smoking may therefore indicate that they understood but did not agree with their desires in a craving state.

In both experiments, it is possible that those participants who revised their choices from one state to the other were mispredicting their preferences at the time of consumption. To conclude this, it would have been necessary that these participants rejected the use of an option that would have committed them to their advance choices. However, because no commitment device was ever introduced, participants were deprived of the possibility to prove their sophistication. In fact, it remains unclear how many of the participants believed that their advance choice was binding and would have preferred it to be binding, because they thought that their preferences at the time of decision would prevail. Maybe (some of) the participants in the experiments were not sharing the authors' Aristotelian assumption that the self that will consume in a different state is the same self that is deciding now and lacked sympathy with this later self.

The sole purpose of the concept of state-dependent preferences lies in the possibility that people's ranking of the causal consequences of their choices might change. Many studies have shown that this is indeed the case and that changes happen often in a very systematic and predictable manner. However, the expectation that a person with a certain ranking of her choices is willing to ignore this ranking for a different one at the time of consumption does not strike us as self-evident. At the very least, this is an assumption that should be empirically validated before declaring state-dependent behavior a manifestation of biased beliefs about oneself.

The problem is that the Aristotelian assumption about state-dependent preferences is notoriously difficult to address empirically. To be able to observe whether people with state-dependent preferences honor their preferences at the time of the decision or those at the time of the consumption requires keeping state-dependent and present-biased preferences apart. Otherwise, a person's deliberate choice in favor of her preferences at the time of the decision over those at the time of consumption could always be an instance of sophistication in the context of present-biased preferences.

Conceptually, the distinction between state-dependent and present-biased preferences is not very difficult to make, but empirically it is. Present-biased preferences operate on a movement along the timeline, and state-dependent preferences on a change in states. However, state changes always come also with a movement along the timeline. While this is a nomological necessity, the passage of time is conceptually irrelevant for state-dependent preferences. Therefore, to test the underlying

assumption about state-dependent preferences, it must be ensured that a possible imposition of current preferences on a later self is indeed due to a disagreement between states and not to the mere passage of time. The only study that tried to do this is Krügel and Uhl (2023). In an experiment on getting up in the early morning, they found a disagreement about the value of early-morning sleep between a rested and an unrested self. This disagreement cannot be attributed to the mere passage of time in their experiment and suggests an intrapersonal conflict in the context of state-dependent preferences. More studies of this type in other decision situations are warranted, however.

The difficulty of studying empirically the underlying assumption of human decision-making in the case of state-dependent preferences also illustrates the difficulty of attributing real examples of dynamic inconsistency to the one source or the other in preference change. State-dependent and present-biased preferences may easily coexist. Thus, a person's ranking of choices may vary with the underlying state, and on top of this, the person may be exposed to present bias as consumption nears. For an impartial, benevolent advisor who observes cases of dynamic inconsistency, it is therefore hardly possible to propose the "right" paternalistic measures.

5 Mispredicting the future and forgetting the past: projection bias and the depletion of rationality

In this section, we take the observation of state-dependent behavior for granted and critically assess the idea of projection bias on a conceptual level. Notice again that the diagnosis of a projection bias is based on the assumption that people view themselves as unitary decision makers who are sympathetic towards changes in their own preferences. In cases of state-induced preference changes, the "true" preferences appear at the time of consumption and it is these that have to be satisfied. Accordingly, advance choices that are systematically influenced by current preferences are a manifestation of systematic mispredictions—the projection bias. Inherent to this concept is the necessary condition that people are unaware of their current misprediction. If they knew that they would choose differently at the time of consumption, the supposedly misguided advance choice would not be the result of a misprediction of preferences. Instead, we would have to relinquish the Aristotelian view of unitary decision makers in this context of state-dependent preferences, because the choosing self apparently imposes its current preferences on the future self.

We intend not to argue here that projection bias does not exist at all but mainly to question its scope. Projection bias is usually thought to explain a wide range of phenomena, from everyday decision-making to suicide (see, e.g., Loewenstein et al. 2003). The potential scope of projection bias is explicitly not limited to once- or twice-in-a-lifetime experiences for which mispredictions may occur simply because of lacking experience. However, the projection-bias explanation of state-dependent behavior is less convincing in daily or recurrent decision situations, precisely because unawareness of the current misprediction is a necessary condition of the concept. In recurrent situations, people do not even have to remember *how* they felt in a certain state in the past to become aware of their bias. They only need to realize

that they previously did not like the consequence of the same choice they are about to make again.

For instance, suppose a person orders a warm winter jacket on a cold day because she overestimates its future value and then returns the jacket upon delivery because the temperature increased and the person is no longer overestimating the value of the jacket, as suggested in Conlin et al. (2007). So, the person orders the jacket because it keeps her warm on cold days and returns the jacket for exactly the same reason. Could this person unconsciously overvalue and order winter clothing on another cold day just to realize a few days later that this was a mistake once again? Notice that it is not necessary for the person to understand how weather affects her valuation of winter clothing. She only needs to remember that she was split over the value of warm clothing in the past to realize that something is wrong. As soon as this happens, the unconscious misprediction becomes a conscious disagreement with her future self regarding the value of warm clothing. In our view, this certainly is conceivable, as the essence of individual decision-making is the weighing of trade-offs. Sometimes the personal ranking of goods goes in one direction, sometimes in another, and every now and then these rankings may be in conflict with each other.

For a decision situation to be recurrent, it is not necessary that iterative states be identical. A decision situation is recurrent if states at two different points in time are more similar to each other than they are to another state regarding the relevant criterion. For instance, suppose a person books her summer vacation trip during the winter and due to projection bias chooses an overly warm destination (the example is borrowed from Loewenstein et al. 2003)—a choice that the person regrets during the summer. The situation becomes recurrent if this person wants to book a summer vacation trip during the following winter again. It seems highly implausible that the person will remain completely unaware of her last summer's misery if she plans to choose an overly warm destination a second time. The corresponding states during summers do not have to be identical. It is sufficient that it is generally warmer in summer and she might get tired of the heat again, whatever the actual temperature will be in the upcoming summer.

The lack of learning in recurrent situations, which is necessary for the concept of projection bias, should not be mixed up with the apparent lack of learning in the context of present-biased preferences. Overvaluing immediate consumption or rewards appears to be largely innate and difficult to break off, even if people are aware of their inclination. Learning in the context of present-biased preferences typically refers to issues of self-control. Projection bias, however, is based entirely on mistaken beliefs. In recurrent situations, a projection-biased person not only holds false beliefs about herself in a future state but must additionally remain unaware of her repetitive mistake. As another illustrative example, suppose a person who is currently satiated falsely believes that she will prefer a salad to a meat at the next day's lunch. Therefore, she preorders the salad and finds out on the next day that she would have liked the meat for lunch. If she again preorders the salad instead of the meat on another day, the concept of projection bias requires her to have forgotten her previous disappointment and to mispredict her desires during lunch once again. Importantly, according to the concept of projection bias, she preorders the salad because she predicts and believes at the time that this is what she wants for lunch on

the next day, not what she thinks on the day of choosing she should have for lunch the day after.

Overall, the projection-bias interpretation of state-dependent behavior requires a degree of forgetfulness about (dis)satisfaction with past choices in recurrent situations that appears implausible for an otherwise fully rational person. At the very least, the assumption regarding forgetfulness should be an explicit part of the concept. In recurrent situations, it is not sufficient that people systematically mispredict their preferences in consumption-relevant states; they must also be forgetful with respect to past choices and corresponding experiences. This certainly raises the question of what remains of rationality if people act “as if” they choose the utility-maximizing consumption bundle, but unfortunately for the wrong set of preferences and with complete absentmindedness towards their own experiences.

6 Concluding remarks

About forty years ago, Schelling (1984b) proposed that we as economists admit “not only unidirectional changes [of preferences] over time, but changes back and forth at intervals of years, months, weeks, days, hours, or even minutes, changes that can entail bilateral as well as unilateral strategy” (p. 6). Schelling called these preferences “alternating preferences” and suggested understanding and modeling phenomena such as these as manifestations of multiple selves. Each of these selves is characterized by its own preferences, which may be identical in many respects. In some others, however, preferences may differ fundamentally so that it seems impossible to compare utility differences between these selves or even add up their utility collectively (Schelling 1984b). Schelling concluded that “this phenomenon of rational strategic interaction among alternating preferences is a significant part of most people’s decisions and welfare and cannot be left out of our account of the consumer” (Schelling 1984b, p. 5). Economists, Schelling noted, could of course ignore these phenomena, but if they do not, they should not have major difficulties in understanding decision makers as a collection of multiple selves (Schelling 1984b). After all, he considered the “art of self-management” as an integral part of the *conditio humana* (Schelling 1978).

Several decades and numerous empirical studies later, we know that preferences can indeed fluctuate systematically, even in short time intervals. The phenomenon of alternating preferences or, in today’s terms, recurrent situations of state-dependent preferences has become difficult to ignore. What in the interim seems to be somewhat forgotten is the idea that these preferences may not simply add up and may instead be expressions of different selves with their own rationality. The underlying similarity between Schelling’s (1984b) anecdote about himself as a young boy who saw a movie about Admiral Byrd’s first Antarctic expedition and subsequently decided to toughen himself against the cold² and the observation in Conlin et al.

² The anecdote goes as follows: “As a boy I saw a movie about Admiral Byrd’s first Antarctic expedition and was impressed that as a boy he had gone outdoors in shirtsleeves to toughen himself against the cold. I decided to toughen myself by removing one blanket from my bed. That decision to go to bed one

(2007) that people tend to return cold-weather items in warmer periods is striking. For Schelling (1984b) it was an illustrative example of alternating selves, while for Conlin et al. (2007), it was an instance of suboptimal behavior based on biased predictions of future preferences. As in the experiments by Read and van Leeuwen (1998) and Sayette et al. (2008), Conlin et al. (2007) cannot disentangle whether people ordering cold-weather items during cold weather would disapprove of their own priorities during warm weather or are naïve about this change in priorities. In any case, the market accommodates their current desires (Sugden 2018).

According to Spiegler's (2019) critique of behavioral economics, the interpretation of state-dependent behavior as a misprediction of one's own preferences and the resulting model of projection bias falls into the category of "functional-form style" modelling, where a standard economic model (here: intertemporal utility maximization with state-dependent preferences) is taken and the behavioral phenomenon is captured by an additional parameter. Spiegler (2019) points out that this type of modelling may distort our understanding of the actual underlying problem and urges putting a stronger weight on a "conceptual framework," especially in the study of behavioral economic phenomena. Regarding our context, we fully agree with Spiegler (2019), because the question of which of people's changing preferences they themselves consider to be relevant is of fundamental importance for theorizing as well as providing policy advice.

Moldoveanu and Stevenson (2001) argue that the Aristotelian view prevailing in the social sciences treats persons as unitary decision makers, while the less popular Heraclitean view accepts that inner conflict is at the core of human existence. This latter view focuses on the processes and phenomena by which people resolve their internal conflicts. The concept of a person may be no less fuzzy than the concept of, say, a firm (see also Parfit 1984). Cowen (1991) has noted that particularly influential literature on self-management bifurcates persons into a long-run, rational self and a short-run, impulsive or "irrational" self. Because it analyzes the strategies that the long-run self uses to induce cooperation from the impulsive self, he refers to this literature as the "command view of self-management." Cowen argues that good self-management instead unleashes forces that enable personality growth. Giving up the asymmetry between the two selves represents a right step in the direction of a new view of self-management that deemphasizes control.

The labels "irrationality" and "bias" (see, e.g., Bernheim and Rangel 2007) are important weapons in the new paternalists' campaign against the self-determined citizen. Behavioral paternalists frequently use the notion of multiple selves (e.g., Thaler and Shefrin 1981; Bernheim and Rangel 2004). In fact, they are dedicated Aristotelians in modeling torn persons as unitary decision makers who happen to have a hyperbolic discounting function or other biases. Originally, the Aristotelian view used to emphasize normative individualism and the dictum that

Footnote 2 (continued)

blanket short was made by a warm boy; another boy awoke cold in the night, too cold to go look for a blanket and swearing to return it tomorrow. But the next bedtime it was the warm boy again, dreaming of Antarctica, who got to make the decision, and he always did it again" (Schelling 1984b, p. 8).

you can't argue about taste (Stigler and Becker 1977). This changes dramatically when the view is complemented with the attribution of behavioral biases. The exponential discounting function, for instance, then becomes a person's telos and thus a normative benchmark. However, the behavioral paternalists' conception of rationality may be overly narrow (Rizzo and Whitman 2019). Intrapersonal conflict may not be reducible to myopia whereby an "impulsive self" does not internalize the effect of its selfish behavior on future alter egos. On the contrary, the "impulsive self" may be the only self that understands the true magnificence of a blissful moment of indulgence (Schelling 1984a).

The spirit of the concept of state-dependent preferences emphasizes the normative equality of the different states. People in the hungry state, for instance, have no less authentic preferences than people in a satiated state or vice versa. Their preferences are simply different in both states. Prima facie, this notion makes paternalistic interventions less plausible than the notion of present-biased preferences where immediate consumption is often reduced to a mere impulse. Paternalism may be defended on the grounds that the government represents the interests of "future individuals" who suffer under the problematic decisions of the present self (Le Grand 2018). In this view, the government is internalizing externalities of the present self on its vulnerable future selves. The concept of state-dependent preferences acknowledges the fluctuation of desires. Although this conflict happens on a timeline, time is not the central concept here. Given their diverging interests, each of the opposing selves imposes an externality on the other. Externalities are reciprocal in the sense of Coase and assigning property rights is a political rather than an economic decision. It is only by the additional and empirically unjustified assumption of an ever-present projection bias that the concept of state-dependent preferences becomes receptive to paternalism.

Sugden (2021) critically points to Sunstein and Thaler's cafeteria example, where the customer's food choice depends on the prominence with which items are displayed, which they interpret as evidence for a "bias." He remarks that the phenomenon merely reflects the well-known psychological effect that people are more likely to choose the items to which their attention is most directed. Psychology, so Sugden, cannot tell us how much attention a person should give to each item and thus does not tell us which choices are biased and which are not: "The idea of bias presupposes a concept of truth, but empirical psychology does not provide—and has no need for—a concept of true preference" (Sugden 2021, p. 16). In line with this, the Heraclitean perspective casts doubts on the confidence with which the new paternalists take sides. The bias may be on the part of the patronizing economist rather than the patronized person.

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akerlof GA (1991) : Procrastination and Obedience. *The American Economic Review (Papers and Proceedings)*. 81: 1–19
- Ariely D, Loewenstein G (2006) The heat of the moment: the effect of sexual arousal on sexual decision making. *J Behav Decis Mak* 19:87–98
- Bernheim BD, Rangel A (2004) Addiction and cue-triggered decision processes. *Am Econ Rev* 94:1558–1590
- Bernheim BD, Rangel A (2007) Toward choice-theoretic foundations for behavioral welfare economics. *Am Econ Rev* 97:464–470
- Buchheim L, Kolaska T (2017) Weather and the psychology of purchasing outdoor movie tickets. *Manage Sci*. 63(11):3718–3738
- Busse MR, Pope DG, Pope JC, Silva-Risso J (2014) The psychological effect of weather on car purchases. *Q J Econ* 130:371–414
- Conlin M, O'Donoghue T, Vogelsang TJ (2007) Projection bias in catalog orders. *Am Econ Rev* 97:1217–1249
- Cowen T (1991) Self-constraint versus self-liberation. *Ethics* 101:360–373
- Jamison J, Wegener J (2010) Multiple selves in intertemporal choice. *J Econ Psychol* 31:832–839
- Krügel S, Uhl M (2023) Is only one of my selves authentic? An empirical approach. *J Behav Experimental Econ* 102:1–9
- Laibson D (1997) Golden eggs and hyperbolic discounting. *Q J Econ* 112:443–477
- Le Grand J (2018) Future imperfect: behavioral economics and government paternalism. *Rev Behav Econ* 5(3–4):281–290
- Loewenstein G (1996) Out of control: visceral influences on behavior. *Organ Behav Hum Decis Process* 65:272–292
- Loewenstein G (2000) Emotions in economic theory and economic behavior. *Am Econ Rev* 90:426–432
- Loewenstein G, Schkade D (1999) Wouldn't it be nice? Predicting future feelings. In: Kahnemann BD, Diener E, Schwarz N (eds) *Well-being: the foundations of hedonic psychology*. Russell Sage Foundation Press, New York, pp 85–105
- Loewenstein G, O'Donoghue T, Rabin M (2003) Projection bias in predicting future utility. *Q J Econ* 118:1209–1248
- Moldoveanu M, Stevenson H (2001) The self as a problem: the intra-personal coordination of conflicting desires. *J Socio-Econ* 30:295–330
- O'Donoghue T, Rabin M (1999) Doing it now or later. *Am Econ Rev* 89:103–124
- O'Donoghue T, Rabin M (2001) Choice and procrastination. *Q J Econ* 116:121–160
- Parfit D (1984) *Reasons and persons*. OUP Oxford, Oxford
- Phelps ES, Pollak RA (1968) On second-best national saving and game-equilibrium growth. *Rev Econ Stud*. 35(2):185–199
- Pollak RA (1968) Consistent planning. *Rev Econ Stud* 35:201–208
- Read D, Van Leeuwen B (1998) Predicting hunger: the effects of appetite and delay on choice. *Organ Behav Hum Decis Process* 76:189–205
- Read D, Loewenstein G (1999) Enduring pain for money: decisions based on the perception and memory of pain. *J Behav Decis Mak* 12:1–17

- Rizzo MJ, Whitman G (2019) Escaping paternalism: rationality, behavioral economics, and public policy. Cambridge University Press, Cambridge
- Samuelson PA (1938) A note on the pure theory of consumer's behaviour. *Economica* 5:61–71
- Samuelson PA (1948) Consumption theory in terms of revealed preference. *Economica* 15:243–253
- Sayette MA, Loewenstein G, Griffin KM, Black JJ (2008) Exploring the cold-to-hot empathy gap in smokers. *Psychol Sci* 19:926–932
- Schelling TC (1978) Egonomics, or the art of self-management. *Am Econ Rev* 68:290–294
- Schelling TC (1984a) Choice and consequence. Harvard University Press, Harvard
- Schelling TC (1984b) Self-command in practice, in policy, and in a theory of rational choice. *Am Econ Rev* 74:1–11
- Simonsohn U (2010) Weather to go to college. *Econ J* 120:270–280
- Spiegler R (2019) Behavioral economics and the atheoretical style. *Am Econ J Microecon* 11:173–194
- Stigler GJ, Becker GS (1977) De gustibus non est disputandum. *Am Econ Rev* 67:76–90
- Strotz RH (1955) Myopia dynamic and utility inconsistency in maximization. *Rev Econ Stud* 23(3):165–180
- Sugden R (2018) The community of advantage: a behavioural economist's defence of the market. Oxford University Press, Oxford
- Sugden R (2021) Normative economics without preferences. *Int Rev Econ* 68(1):5–19
- Thaler RH, Shefrin HM (1981) An economic theory of self-control. *J Polit Econ* 89:392–406
- von Hayek FA (1989) The pretence of knowledge. *Am Econ Rev* 79:3–7

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.